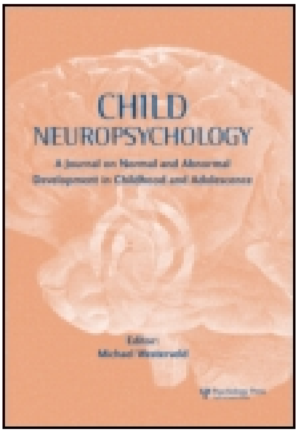
Serial and subjective clustering on a verbal learning test (VLT) in children aged 5–15: The nature of subjective clustering

DOI:
10.1080/09297049.2012.670215

Document status and date:
Published: 01/01/2013

Document Version:
Peer reviewed version



Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ncny20

Serial and subjective clustering on a verbal learning test (VLT) in children aged 5-15: The nature of subjective clustering

Celeste Meijs [a] , Petra Hurks [b] , Nico Rozendaal [c] & Jelle Jolles [d]
[a] Ruud de Moor Centre , Open University , Heerlen , The Netherlands
[b] Faculty of Psychology and Neuroscience , Maastricht University , Maastricht , The Netherlands
[c] Faculty of Health, Medicine and Life Sciences , Maastricht University , Maastricht , The Netherlands
[d] Faculty of Psychology and Education, VU Amsterdam , Amsterdam , The Netherlands
Published online: 16 Mar 2012.

Child Neuropsychology, 2013
Vol. 19, No. 4, 385–399, http://dx.doi.org/10.1080/09297049.2012.670215

Routledge
Taylor & Francis Group

Serial and subjective clustering on a verbal learning test (VLT) in children aged 5–15: The nature of subjective clustering

Celeste Meijs[1], Petra Hurks[2], Nico Rozendaal[3], and Jelle Jolles[4]

[1]Ruud de Moor Centre, Open University, Heerlen, The Netherlands
[2]Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands
[3]Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands
[4]Faculty of Psychology and Education, VU Amsterdam, Amsterdam, The Netherlands

This study investigated which strategies children aged 5–15 years ($N = 408$) employ while performing a multitrial free recall test of semantically unrelated words. Serial clustering (i.e., a relatively passive strategy) is an index of the sequential consistency of recall order. Subjective clustering (i.e., a more active strategy) is based on similar word groupings in successive trials. Previously, Meijs et al. (2009) found that the level of (serial and subjective) clustering increases with age. At all ages, the level of serial clustering correlates positively with the ability to recall information on VLT trials. However, subjective clustering is more predictive of VLT performance than serial clustering after ≥ 3 trials, but only in children aged 8+. Knowledge on how children organize words (based on, for example, sound or meaning) and how this relates to developmental stage is still lacking. This study revealed that the level of subjective clustering is primarily determined by the position of words in a VLT list. More specifically, primacy (i.e., recall of words 1–3 of the VLT list — whether recalled in the same order or reversed) and recency (i.e., recall of words 14–15) effects primarily determine level subjective organization over successive trials. Thus, older children still organize words based on the serial position of the VLT list and are much less likely to organize them based on any other feature of the words, for example, sound or meaning. This indicates that the most important information to be learned needs to be presented first or last, even in older children and even with repeated presentations.

Keywords: Serial clustering; Subjective clustering; Development.

Over the years, much research has been directed at explicating the structure of semantic memory and, more specifically, the nature and relevance of operations (or organization strategies) that take place when retrieving information from short-term memory or long-term memory. In this context, organizational strategies have generally been conceived as mentally, effortful processes that are implemented deliberately by the individual in order to enhance their memory performance (Bjorklund & Douglas, 1997). Indeed, an efficient use of memory organization strategies has been found to improve verbal learning

Address correspondence to Celeste Meijs, Open University, Ruud de Moor Centre, Heerlen, The Netherlands. E-mail: celeste.meijs@ou.nl

performance in both children and adults (Bjorklund & Douglas, 1997; Bukatko & Daehler, 1998; DeMarie & Ferron, 2003; Meijs et al., 2009; Ornstein, 1999; Schneider, Knopf, & Stefanek, 2002; Siegler & Booth, 2004). In contrast, less efficient use of organization strategies have been found to cause a lower performance on verbal learning tests (Gaultney, Bjorklund, & Goldstein, 1996). For example, individuals with frontotemporal dementia performed significantly less well on a free recall test, presumably because they failed to implement sophisticated organizational strategies during learning (Glosser, Gallo, Clark, & Grossman, 2002). Finally, several researchers have claimed to have found increased verbal memory performance after explicit (i.e., on a metacognitive level) training of strategy use (e.g., Kofman, Larson, & Mostofsky, 2008; Schwenck, Bjorklund, & Schneider, 2009). However, the exploration of the potential of organizational memory strategies in pediatric and developmental neuropsychological research has only begun recently (e.g., Meijs et al., 2009). This is unfortunate, since insights in how children use strategies may help in adapting the presentation of information (e.g., at school) to children in an age-appropriate manner, so they can most optimally benefit from strategy use according to their abilities.

To gain more insights, Meijs et al. (2009) previously studied the developmental trajectories of two types of verbal memory strategies, namely serial clustering and subjective clustering, during a multitrial Verbal Learning Test (VLT) that consisted of 15 semantically unrelated words. In serial clustering, words are recalled in the same order as presented; that is, the information is not actively handled or rearranged by the child. This strategy is considered to be a relatively simple, passive strategy. Results from this study indicated that children frequenting kindergarten (i.e., $M_{age} = 6.51$ years) already made use of serial clustering. This is in line with studies reporting that children as young as 4 years are able to use (simple) strategies in general (Cox, Ornstein, Naus, Maxfield, & Zimler, 1989; Schwenck et al., 2009). In addition, Meijs and colleagues (2009) found that, in their population of children aged 6–12 years, the level of serial clustering remained stable with age.

The second strategy included in Meijs' study was subjective clustering. Subjective clustering is defined in terms of the subject's tendency to recall items in the same order on different trials in the absence of any experimentally manipulated sequential organization among items in the stimulus list (in line with Tulving, 1962). This type of clustering enhances recall, because if one word of the cluster is recalled, the other word is more readily recalled as well. This type of clustering is broader than — the usually studied— semantic clustering, in which the words are coupled based on the semantic category they belong to and/or semantic relationships they have (for more information on semantic clustering see, for example, Bjorklund & Douglas, 1997; Cole, Frankel, & Sharp, 1971; Kramer, Delis, Kaplan, O'Donnel, & Prifitera, 1997; Schlagmuller & Schneider, 2002; Siegler & Alibali, 2005). Because subjective coupling is based on all possible kinds of connections between the words (e.g., similar sounds, experiences of children that connect some words, and mental images that are generated that contain more than one word), semantic clustering is only one subtype of clustering that can be measured by subjective clustering. The word list used in this study contained only semantically unrelated words, thereby limiting the possibility of taking advantage of "standard" semantic relationships. If words were coupled, this was based on a connection between the words that was logical to the children.

The previous study of Meijs and colleagues (2009) showed that children from age 10 (i.e., Grade 4) onwards made more use of the above-mentioned subjective clustering than younger children. Also, subjective clustering was superior in this older group when information was presented repeatedly, that is, after three or more trials. However, content knowledge on how these older children organize words individually is still lacking here. The question remained whether older children actually coupled words based on more elaborated strategies, such as similar sounds, events, or stories that made the words belong together in the child's opinion, were based on the formation of mental images including more words, or that the couplings took place on a less sophisticated level (e.g., the previous discussed serial clustering). Therefore, by the use of new methods of analyzing word retrieval in VLT, we investigated in the present study what kind of connections (e.g., similar sounds) were used more specifically by children of varying ages and whether the connections between the words became more sophisticated with age.

It is imperative to study the base of the connections between words because this gives insight into the development of information processing. Findings from the present study can be used to give age-appropriate recommendations for an optimal presentation of and/or successful interventions aiming to increase the capacity to learn information. For instance, if older children are able to make more sophisticated connections between the words, this makes it easier for them to remember unrelated information as compared to younger children.

METHOD

Participants

The present study was part of a large cross-sectional and longitudinal study on the cognitive development of school-aged children (for more detailed information, see Meijs, 2008). Only cross-sectional data were used in the present study. Participants were recruited from 29 regular primary and secondary schools in the city of Maastricht and surroundings (The Netherlands). The parents/caregivers (referred to as *caregivers* hereafter) of the children enrolled in kindergarten, Grades 2, 4, 6 (primary schools), and 7 and 8 (secondary schools; ranging from lower secondary professional education to preuniversity education) received an information package via the school describing the purpose of the study, a request to participate, a questionnaire, a form to give consent for the child to participate, and a stamped-addressed envelope.

Of the 1086 caregivers who replied, 892 (82%) gave consent for their child to participate. Children who met the following criteria were eligible for participation: The child (a) had not repeated and/or skipped a grade (indication for "normal' development), (b) had the Dutch nationality (we assumed that these children primarily used Dutch as their primary language) and (c) did not use medication that is known to influence cognitive performance (for instance, asthma medications). In total, 431 children (215 boys, 216 girls) participated in this study (for a more detailed description of the demographics of the participating children, see section on statistics and missing data as well as Table 1).

All children completed a battery of neuropsychological tests, including tests measuring general verbal ability (VIQ; see Table 1), memory, language comprehension, and time estimation. Well-trained psychological assistants (i.e., undergraduate psychology students) tested the children in a room at the schools. Testing took approximately 1.5 hours; all tests were administered in the same order for each child. The Ethics Committee of the Faculty of Psychology of Maastricht University approved the study protocol.

Table 1 Demographics of the Children.

Grade	Age (mean [*SD*])	Sex		VIQ
		Boys	Girls	
K	6.31 (0.33)	34	35	11.08
2	8.37 (0.44)	36	31	10.84
4	10.35 (0.36)	33	45	10.40
6	12.30 (0.41)	31	35	9.97
7	13.37 (0.38)	39	31	9.94
8	14.44 (0.33)	31	27	9.77
$N = 408$		204	204	$M = 10.34$, $SD = 2.48$

Note. K = kindergarten.

Instruments

VLT. A verbal learning test (VLT) was administered to measure verbal learning and strategy use. A VLT is often used in clinical settings and memory research and is one of the most sensitive verbal memory tests. Its test-retest reliability is reported to be high (Lezak, 1995; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2005). The test used in the current study was based on the Rey Auditory Verbal Learning Test (RAVLT; Schmidt, 1996) but differed in the words used. The words used in the VLT test have been checked for their frequency of use (Linschoten, 1963) and imageability (Van Loon-Vervoorn, 1989). The procedure consisted of five trials of presentations of 15 familiar objects. They were presented to the children in the same order in each trial (i.e., immediate recall condition). After each presentation of the items, the children had to verbally recall as many items as they could. After about 15 to 20 minutes, during which no memory interfering tests were administered, the children were asked to recall as many items as possible from the list, without prompting (i.e., delayed recall condition). This was followed by a recognition trial, in which 30 items were presented, of which 15 items were from the immediate recall task and 15 were new items. The children had to answer "Yes" or "No" to whether the item belonged to the list of items presented to them in the immediate recall condition. The current study was part of a longitudinal study into cognitive development; therefore, three statistically parallel versions of the test were used. Additionally, the words were presented as pictures to half of the children and auditory to the other half to study the effects of presentation modality on recall (for more information, see Meijs, 2008). The children were divided evenly over the different versions and presentation modalities. The results of the clustering of words were not influenced by the factor "version." The order of the recall of the words was documented and used for the calculation of serial and subjective clustering.

Serial Clustering. Serial clustering was used to quantify simple, passive strategy use. It was calculated by dividing the number of observed serial clusters by the number of serial clusters that could be expected by chance (for more information, see Meijs et al., 2009). The number of observed serial clusters equals the number of times that two items that were presented sequentially were recalled sequentially. The number of chance serial clusters in a trial was the maximum number of possible serial clusters divided by the sum of the number of items recalled correctly, plus the number of repetitions and intrusions (i.e., items that did not belong to the presented items) on that trial. The maximum number of possible serial clusters was calculated as follows: All items were given a ranking (1 to 15) according to the place of the item in the presentation order. Intrusions had no ranking, because they were not in the presentation list and repetitions had the same ranking as the first recall of that particular item. Then, the recalled items were placed in ascending order according to their ranking. If there were adjacent numbers in the ranking, this was a possible serial cluster. The number of possible item clusters was divided by the total number of items recalled correctly, plus the number of repetitions and intrusions. For example, if Item 5, Item 7, Item 8, Item 2, Item 15, and Item 1 were recalled (in that order) and placed according to their ranking in the order 1 – 2 – 5 – 7 – 8 – 15, then two possible item clusters were recalled, namely, Items 1 and 2, and Items 7 and 8. Thus, the maximum number of possible serial clusters in this example was two; and the total number of words recalled was 6. The number of chance clusters in this example is $2 / 6 = 0.33$. The number of observed clusters in this example is 1 (Items 7 and 8 in the recall order). This leads to a serial clustering of $1 / 0.33 = 3.03$. By means of this calculation, a correction for total number of words recalled is included. This calculation is a modification of the calculation described by Saan and Deelman (1996), in that we included intrusions in the calculation, in order to limit “noise” in the serial-clustering calculation (Saan & Deelman, 1996). The “noise” can be caused by the false detection of serial clusters because of the omission of intrusive items between otherwise adjacent words. There were five serial-clustering calculations, one for each trial of the VLT. In the present study, we made use of a summed score of serial clustering over trials.

Subjective Clustering. Subjective clustering was used to study complex, organizational strategy use. It was calculated as follows (for more information, see Meijs et al., 2009):

$$\mathrm{SC} = \mathrm{ITR} - \left(\frac{2\mathrm{C}\,(\mathrm{C}-1)}{\mathrm{H}\times\mathrm{K}}\right). \tag{1}$$

In this calculation, (a) SC stands for Subjective Clustering, (b) ITR stands for the number of Intertrial Repetitions (i.e., the number of groups of two clustered items that were recalled on Trial t and the subsequent Trial $t+1$), (c) C stands for the number of items recalled on Trial t but also on Trial $t+1$, (d) H stands for the sum of the number of items recalled correctly, repetitions, and intrusions on Trial t, and (e) K stands for the sum of the number of items recalled correctly, repetitions, and intrusions on Trial $t+1$. The calculation was a modification of the calculation described by Saan and Deelman (1996) and Sternberg and Tulving (1977), in that we included intrusions in the calculation in order to limit “noise” in the subjective clustering calculation (Saan & Deelman, 1996; Sternberg & Tulving, 1977). ‘Noise’ could be caused by the false detection of subjective clusters if intrusive items were omitted that would lead to the finding that two items were recalled in sequence whereas there was actually an intrusive word between them. There were four subjective-clustering calculations, one for each trial pair (i.e., two successive trials) of the VLT.

However, based on the above-mentioned definitions, a serially clustered group of responses that persists across trials will be scored as both serial clustering but also as a subjective clustering. Therefore, an additional subjective-clustering score was calculated to explore whether this occurred in the present study. This formula extracts all serial clusters (i.e., in the forward direction only, in line with the serial-clustering formula) from the subjective clustering. This calculation of subjective clustering will be referred to as *subjective clustering* (i.e., summed over trials) in the remainder of this article. This latter form of subjective clustering is a more pure form of clustering because the serial component is omitted and will therefore be the focus of the current study (next to serial clustering). Additionally, we studied the exact nature of the coupling of the words.

Metacognition Question into Strategy Use VLT. After the administration of the VLT, the children were asked, amongst others, if they made groups of the words during recall. The children answered this question by saying “Yes” or “No.” This test was comparable to the method used by O’Sullivan (1996), who studied children’s metamemory about the influence of conceptual relations on recall (for more information, see O’Sullivan, 1996).

Verbal Ability (VIQ). As an estimate of verbal ability (VIQ), standard scores of the Vocabulary subtest of the Dutch Wechsler Intelligence Scales Revised (WISC-Rn; De Bruin et al., 1986) were collected. The WISC-Rn, rather than the WISC-III, was used because the latter became available in the Netherlands after the study had started. In this test, the children have to explain words, ranging from easy to complex. The standard score ranges from 1 to 19 ($M = 10$, $SD = 3$). The reliability and validity have been described as average to good (De Bruin et al., 1986).

Statistics

Missing Data and Extreme Values. Before data analyses, we checked for unreliable testing, missing data, and extreme values. First, data from the children of whom the test administration of the VLT was unreliable according to the test administrator or not executed were excluded from the analyses ($n = 23$). Unreliability was due to several factors including technical problems such as failure of equipment, refusal of the child to cooperate, low motivation (for instance, due to fatigue), or missing data due to time constraints. These missing data were not replaced because only approximately 5% of data were missing per measure (Croy & Novins, 2005). Next, we checked for extreme values, defined as values that are minimally three times the interquartile distance above the 75th percentile or below the 25th percentile (Huizingh, 2002). No extreme values were found. Thereby, the final dataset included 408 unique cases (see Table 1).

Serial and Subjective Clustering. First, while correcting for grade and VIQ, partial correlation analyses were performed correlating total score on the VLT on the one hand and serial clustering and subjective clustering (with or without correction for serial clustering) on the other. These analyses were conducted to confirm the hypothesized positive relation between clustering strategy use (independent of type of strategy) and overall recall on the VLT, while correcting for age and VIQ.

Secondly, to study whether there were differential developmental patterns for subjective clustering with and without correction for serial clustering, a General Linear Model (GLM) Repeated Measures analysis was performed with subjective clustering with correction for serial clustering and subjective clustering without correction for serial clustering as dependent factors, grade as independent factor, and VIQ as a covariate.

Next, we wanted to study the preferences of children (i.e., the use of serial clustering versus subjective clustering) and whether there were differential developmental patterns for serial and subjective clustering. Therefore, we decided to focus on subjective clustering with correction for serial clustering only. In this context, a GLM Repeated Measures analysis was performed with serial and subjective clustering (corrected for serial clustering) as dependent factors, grade as independent factor, and VIQ as covariate.

Next, in case significant effects for grade were found on the repeated measures analyses, serial and subjective clustering (corrected for serial clustering) were analyzed by use of two separate GLM univariate analyses to study exactly where there were differences in the developmental patterns. Serial and subjective clusterings summed over trials were defined in the model as dependent factors, grade was the independent variable, and VIQ was included as covariate.

In addition, the nature of the coupling of the words was analyzed, independent of the formulas for subjective and serial clustering defined above. More precisely, the most common clusters of words were produced by a formula that counted the times that two words were recalled together on one trial and the next. These results were plotted as a function of word frequency. This was followed by further analyzing the word clusters with the highest frequencies. To verify the patterns shown by the plots, GLM univariate analyses with clusters of the words that had the highest frequencies (namely words in first and last parts of the list: clusters of Words 1 and 2, 2 and 3, 1 and 3, 13 and 14, 13 and 15, and 14 and 15) as dependent variables and grade as dependent variable (average frequency of the cluster per trial pair) were executed. These last analyses made it possible to study clustering of words that were serially linked in backward order, restricted to the words of interest, namely the first and last words of the word list.

Whether or not children reported to have made groups of the words during recall (Metacognition Question into strategy use VLT) was scored as “Yes” or “No.” Chi-square analyses were performed to study grade effects, and independent-sample *t*-tests with serial and subjective clustering (without serial clustering) as dependent factors and reported use of grouping of words during the recall as independent factor were performed to study whether children that reported to have grouped the words actually had a higher clustering.

Verbal ability, or verbal IQ, was included in the analyses as a correcting factor, given the fact that this factor has been reported to influence verbal learning and strategy use (Berk, 1998; Schneider, Kron-Sperl, & Hunnerkopf, 2009; Vakil, Greenstein, & Blachstein, 2010). Sex was not included in the analyses because a previous study showed no effects of sex on clustering (Meijs, 2008).

RESULTS

Partial Correlations VLT, Subjective and Serial Clustering

Significant correlations were found for (a) total score on the VLT and serial clustering ($r = .28$, $p \leq .001$, $df = 372$), (b) subjective clustering without serial clustering ($r = .27$, $p \leq .001$, $df = 372$), and (c) subjective clustering with serial clustering ($r = .47$, $p \leq .001$, $df = 372$).

Developmental Patterns of Serial and Subjective Clustering With or Without Serial Clustering

Significant interaction effects for Grade x Type of subjective clustering were found, Pillai’s trace: $F(1, 371) = 10.21$, $p = .002$, $\eta_p^2 = .03$. Figure 1 reveals that subjective

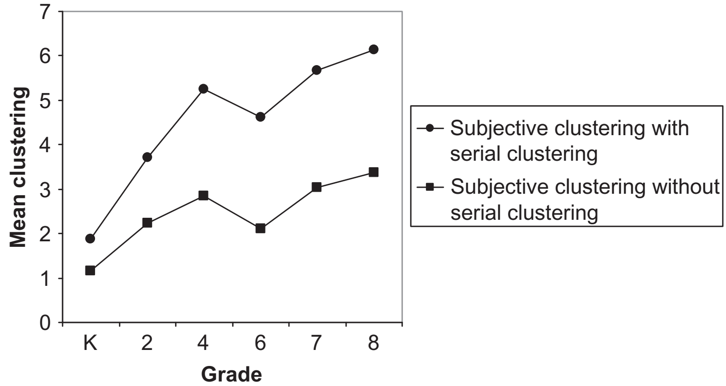


Figure 1 Developmental patterns of subjective clustering with and without serial clustering by grade, corrected for VIQ.

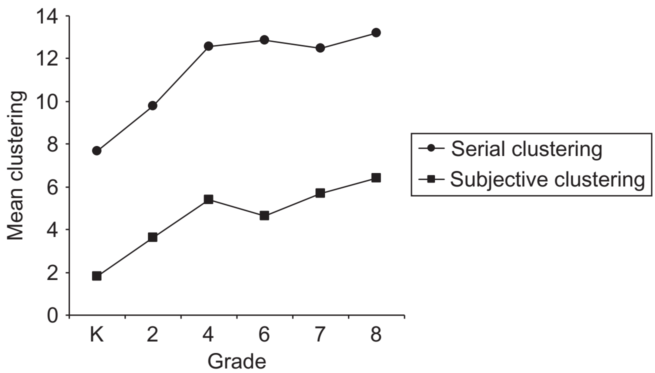


Figure 2 Developmental patterns of serial and subjective clustering by grade, corrected for VIQ.

clustering with serial clustering inclines faster in the younger grades than clustering without serial clustering.

Significant interaction effects for Grade x Type of clustering (serial clustering and subjective clustering with correction of serial clustering) were found, Pillai's trace: $F(1, 369) = 30.81$, $p < .001$, $\eta_p^2 = .08$. Figure 2 shows that serial clustering inclines faster in the younger grades than subjective clustering.

Serial Clustering

Significant effects were found for grade, $F(5, 369) = 6.89$, $p \leq .001$, $\eta_p^2 = .09$. Children in kindergarten made less use of serial clustering than children in Grades 4+ (see Figure 2 and Table 2). However, even though children in kindergarten made less use of serial clustering than older children, on average, they were able to use it based on the finding that the calculation has an outcome above 0 (i.e., $M = 7.67$, $SD = 0.84$; see Figure 2).

Subjective Clustering Without Serial Clustering

Significant effects were found for grade, $F(5, 371) = 6.04$, $p \leq .001$, $\eta_p^2 = .08$. Children in kindergarten made less use of this type of subjective clustering than children in

Table 2 Post Hoc Pair-Wise Comparisons per Grade Couple for Serial Clustering, Subjective Clustering, Word Couple 1–2, Word Couple 2–3, and Word Couple 14–15.

Grade	Couple	Serial Clustering	Subjective Clustering	Word Couple 1–2	Word Couple 2–3	Word Couple 14–15
K	2	.991	.354	1.000	.197	1.000
	4	.000	.001	.014	.003	1.000
	6	.000	.573	.586	.099	1.000
	7	.001	.001	.020	.049	.072
	8	.000	.000	.024	.078	1.000
2	4	.194	1.000	.812	1.000	1.000
	6	.110	1.000	1.000	1.000	1.000
	7	.269	.968	1.000	1.000	.083
	8	.082	.163	1.000	1.000	1.000
4	6	1.000	.848	1.000	1.000	1.000
	7	1.000	1.000	1.000	1.000	1.000
	8	1.000	1.000	1.000	1.000	1.000
6	7	1.000	.620	1.000	1.000	1.000
	8	1.000	1.000	1.000	1.000	1.000
7	8	1.000	1.000	1.000	1.000	1.000

Note. Adjustment for multiple corrections: Bonferroni.

Grades 4, 7, and 8 (see Figure 1 and Table 2). Even though younger children made less use of subjective clustering than older children, this group was able to use it based on the finding that the calculation has an outcome above 0 (i.e., $M = 1.18$, $SD = 0.34$; see Figure 1). Children in Grade 6 showed a "dip" in subjective-clustering use. We are not able to explain these findings, more so because they were not in line with our expectations based on our previous study in which we made use of another dataset (Meijs, et al. 2009).

Frequency of Clusters Made on VLT

Exploration of the frequency of clusters plot (see Figure 3) shows that the words in the first and last parts of the list were clustered more often than the words in the middle part of the list (grade was not considered in this figure). The plots were made to select the word couples for further analyses. Analyses of all word couples would lead to an enormous amount of analyses, which in turn could lead to more Type II errors. Further exploration of inspection of the frequency of clusters plots showed that the words in the first part of the list (Words 1-2-3) and the words in the last part of the list (Words 13-14-15) were most often recalled together on one trial and the next. Figure 4 shows the distribution of the frequency of Words 1-2-3 and 13-14-15 per grade. These word couples were analyzed further and showed significant grade differences for word couple 1–2, $F(5, 18) = 4.74$, $p = .060$, $\eta_p^2 = .57$, word couple 2–3, $F(5, 18) = 4.74$, $p = .006$, $\eta_p^2 = .57$, and word couple 14–15, $F(5, 18) = 2.87$, $p = .045$, *PES* $\eta_p^2 = .44$. Note that this pattern was found for all three parallel versions of the test, which was a solid clue for position effects over clustering based on features of the words.

Metacognition Question into Strategy Use VLT

In total, 108 children reported that they grouped the words during recall — independent of content — and 284 children reported that they had not grouped the words during recall. No significant effect of grade was found ($X^2 = 3.48$, $df = 5$, $p = .626$). This

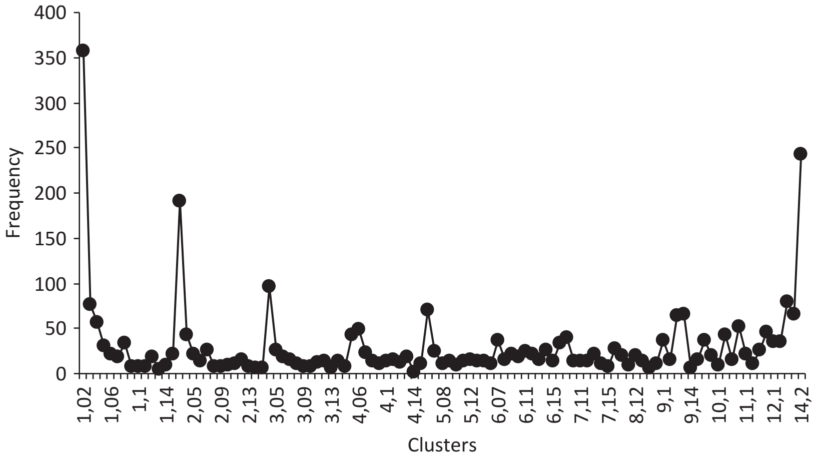


Figure 3 Frequencies of clusters per cluster.

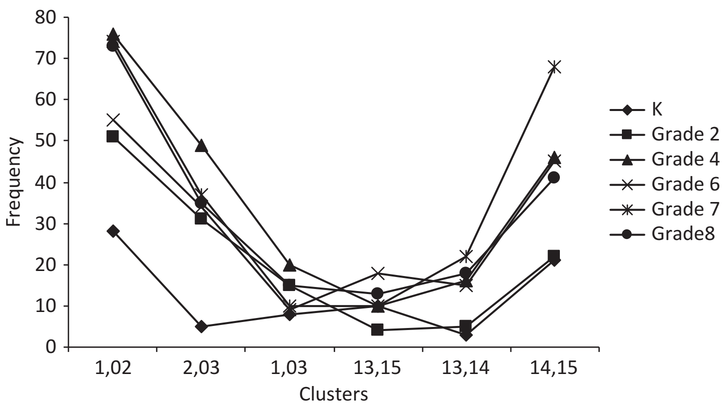


Figure 4 Frequencies of clusters for words 1, 2, 3, 13, 14, and 15 per grade.

indicates that there was no increase in reports of having grouped the words with age. There were no differences in serial and subjective clustering (while correcting for serial clustering) between children that reported to group the words and children that reported not to have grouped the words (respectively, $t = 1.53$, $df = 389$, $p = .176$ and $t = 1.52$, $df = 390$, $p = .688$).

DISCUSSION

The development of serial and subjective clustering on the trials of a verbal learning test (VLT) was studied in normally developing children ages 5 to 15. In short, serial clustering (i.e., a relatively passive strategy) is an index of the sequential consistency of recall order. Subjective clustering (i.e., a more active strategy) is based on similar word groupings in successive trials. As mentioned in the method section, if used consistently, serial clustering could be considered as a type of subjective clustering. However, because we wanted to compare serial clustering to "other" types of clustering, we corrected in our "subjective clustering" formula for serial clustering use.

First, we established that strategy use (independent of type of strategy used) correlated with better performance on the VLT. Thus, the children, that were able to use (both) the strategies, obtained the highest recall scores on the VLT. This is line with findings reported in the literature (Cole et al., 1971). Correlations did not significantly differ as a function of type of clustering. This subscribes the relevance of studying both types of strategies in VLT performance. In line with this, we found a differential developmental pattern of serial and subjective clustering with age (see Figure 2). Serial clustering inclines faster in the younger grades than subjective clustering. These results are in line with Meijs et al. (2009). To study this pattern in more depth, separate analyses were performed into differences in strategy use with age. Note that both forms of clustering remain in use as children grow older. Apparently, even though refinement of strategies and improved decision-making abilities continues during late childhood and early adolescence (P. Anderson, 2002), the older participants tended to hold on to cautious and more conservative strategies (i.e., serial clustering) as well while performing a VLT task.

Based on the findings of Meijs et al. (2009), we expected that serial clustering would be used from kindergarten onwards and that its use would remain stable with age. However, the findings from the present study indicated that this was partially true; that is, serial clustering summed over trials increased with age and became stable after Grade 4 onwards. We hypothesize that this difference in findings is the result of increased statistical power and a wider age range. Meijs et al.'s (2009) study included 79 children in the Grades K through Grade 6, whereas the latter study included 408 children in the Grades K through Grade 8. Thus, even though all children are able to use serial clustering, there is an increase in use up until Grade 4 (i.e., age 10) and a stabilization after that age. This is in line with studies reporting stabilization of performance on a VLT around the age of 10 to 12 (Korkman, Kemp, & Kirk, 2001).

For subjective clustering, we expected based on the findings of Meijs et al. (2009) that children in Grade 4 (age 10) would make more use of subjective clustering than younger children. This is indeed what the majority of our results showed; that is, the findings of the present study suggest an increase in subjective clustering with age up until at least Grade 8 (mean age 14.44 years). This seems to be in contrast to, for example, V. Anderson et al.'s (2001) study on the Rey-Osterriech Complex Figure Test (i.e., a visuospatial memory task), which revealed that, despite access to a greater repertoire of strategies, conceptual strategies to (more simple) piecemeal strategies seemed to occur around 12–13 years of age (V. Anderson, Anderson, Northam, Jacobs, & Catroppa, 2001). However, these authors claim that these fragmented strategies can be just as effective if planned carefully and logically. Differences between studies (even while studying only word-learning tests) complicate direct comparisons. Still, the results of the present study seem to suggest that developmental changes in subjective organizational strategies occur until at least middle adolescence.

In the present study, we focussed specifically on the nature of the coupling of the words during subjective clustering, which is in our opinion the first study into this subject (note that most studies have been using semantically related word lists). The couplings of words can occur on a simple level, such as connections based on similar sounds, or on a more complex level, such as based on imagery (i.e., making a mental picture in the mind that includes representations of the words) or elaboration (i.e., making, for instance, a story that contains the words) or combinations of these. Couplings based on semantic categories were not possible because of the use of a word list with semantically unrelated words. We expected older children to make use of more elaborated couplings between words. However, analyses of the nature of the coupling of the words revealed that subjective clustering was mainly due to clustering of the words based on the position of the words in the presentation list. This is an indication that serial clustering is involved in the incline in subjective clustering with age. Older children recalled more words from the first and the last portion of the list (respectively Words 1 and 2, Words 2 and 3, and Words 14 and 15) together in sequential trials than younger children. Thus, it appeared that the words that older children clustered on one trial and the next were words that were presented in succession in the presentation list but in reversed order (note that subjective clustering was corrected for serial clustering). Although we expected older children to make use of a more elaborated clustering of words based on for instance the meaning of the words, the position of the words in the presentation list was the basis of the coupling and is hence a more passive manner of information processing than we expected. However, this form of clustering is still more active than serial clustering because the words are clustered in a backward order. Cole et al. (1971) describe, already in 1971, that older children recall words in an organized manner, whereas for younger children, there is no recognizable organization in output order. They speculate that if younger children would be guided to use strategies, they also would be able to use (organization) strategies, but recent studies into the development of strategy use contradict these speculations. Recent insights into strategy use report that if a strategy is first used, this might not lead to better performance because of the effort that is needed to execute the strategy (for more information, see Schlagmuller & Schneider, 2002).

We can only speculate about the reason why older children do not cluster words on more elaborated manners than younger children on our VLT. Most importantly because word lists that make use of semantically related words (e.g., the California Verbal Learning Test) show that children are able to actively rearrange the words into categories (Kramer et al., 1997). However, in, for example, the CVLT, words from the list belong to four unique semantic categories. Also, children are explicitly asked to recall words from a specific conditions (e.g., which fruits do you recall from the list?). It is possible that, in semantic category tests, children are implicitly stimulated to semantically cluster words instead of word order. This procedure deviates from our VLT.

Another strong indicator for the presentation-list effect in our study was that these findings were found for all three parallel versions of the list used in the present study. This makes the conclusion that it is a presentation-list effect over an effect of the meaning or features of the words even more solid. These findings have important implications for the optimal presentation of information to children. Information that is presented first or last is remembered best. These findings are in line with the conclusions drawn from an extended database of studies on primacy/recency effects in word retrieval, which however primarily focussed on adult populations. In general, words from the first part of a word list (primacy) and the last part of the word list (recency) are remembered better than words in the middle part of a word list (Lezak, Howieson, & Loring, 2004). The relatively few pediatric studies are still inconsistent in their outcome. Cole et al. (1971) describe that age differences in recall (on a word list that contains semantically related words, studied in children ranging from Grade 1 to 9) are the result of differences in the amount of words that can be retrieved form long-term storage. These words are particularly words from the first (primacy) and middle parts of a word list and these words (most striking on Trial 1) show the tendency to be recalled in the same order as they were presented. Recall of the words in last part of the word list (recency items) showed no age effect in that particular study, which is interpreted as that no development of short-term memory takes place (Cole et al., 1971). De Alwis, Myerson, Hersey, and Hale (2009) also reported primacy and recency effects in studies into list learning in children (aged 6 to 16, studied with a multitrial supraspan test), but age effects were only found for primacy words. They ascribed the findings to age-related differences in strategy, but an alternative explanation is a developing secondary memory (which can be compared to long-term memory) as opposed to a not developing primary memory, which they referred to as similar to working memory. The working memory is described to be involved in the recall of recency items (for more information, see De Alwis et al., 2009). In their paper, they refer to the findings of Cowan, Nugent, Elliott, Ponomarev, and Saults (1999) and Cowan, Nugent, Elliott, and Saults (2000), who describe that working memory does improve with age, but they provide no solid explanation for these differences in findings (Cowan et al., 2000; Cowan et al., 1999). Forrester and Geffen (1991) reported serial position effects, but no age differences in primacy and recency scores. These latter authors even state that (v)ariation across age groups was due to the number of words recalled and not to the position the words occupied within the list (Forrester & Geffen, 1991). Thus, there is still no consensus in the literature regarding age differences in primacy and recency effects. However, the results in the present study suggest that they exist for primacy as well as for recency words (i.e., in secondary as well as in primary memory). Studies more focussed on this subject should shed more light on this subject.

Additionally, Cole et al. (1971) report that serial position effects are reduced if semantic organization is used. They interpret this as that these strategies work counter to each other (Cole et al., 1971). This is not the case in the present study, where serial and subjective clustering are both in use, because they found subjective clustering is based on serial position effects.

On a metacognitive level, we explicitly asked the children post hoc if they made groups of the words during recall. No effects of grade and strategy use were found; that is, the same percentage of children in all grades reported to have grouped the words. Thus, we found no age effect on this measure and we found no differences in serial and subjective clustering between children that reported to have grouped the words and children who reported not to have grouped the words. Knowledge about the use of memory strategies is called metamemory. Besides knowledge on what strategies are best to use in a certain situation, the child has to be aware that it actually used a strategy. It is possible that serial and subjective clustering is (for some children) unconscious and, therefore, not reported. Thus, children might have used a strategy but, because it was outside their awareness, they do not recall using it. Our finding is in line with other studies into the relation between memory and metamemory that describe that there was no relation between measures of metamemory and memory performance (O'Sullivan, 1996). Additionally, studies focussing on development trajectories of strategy use describe that, when children start to use a strategy, this is not always effective and might even lead to a lower performance (Schlagmuller & Schneider, 2002). This could explain why there was no relation between reported strategy use and performance.

Original manuscript received September 13, 2010
Revised manuscript accepted February 11, 2012
First published online March 16, 2012

REFERENCES

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, *8*, 71–82.

Anderson, V., Anderson, P., Northam, E., Jacobs, R., & Catroppa, C. (2001). Development of executive functions through late childhood and adolescence in an Australian sample. *Developmental Neuropsychology*, *20*, 385–406.

Berk, L. E. (1998). *Development through the lifespan*. Needham Heights, MA: Allyn & Bacon.

Bjorklund, D. F., & Douglas, R. N. (1997). The development of memory strategies. In N. Cowan & C. Hulme (Eds.), *The development of memory in childhood* (pp. 201–246). Hove, East Sussex, United Kingdom: Psychology Press.

Bukatko, D., & Daehler, M. W. (1998). *Child development: A thematic approach* (3rd ed.). Boston, MA: Houghton Mifflin.

Cole, M., Frankel, F., & Sharp, D. (1971). Development of free recall in children. *Developmental Psychology*, *4*, 109–123.

Cowan, N., Nugent, L. D., Elliot, E. M., Ponomarev, I., & Saults, J. S. (1999). The role of attention in the development of short-term memory: Age differences in the verbal span of apprehension. *Child Development*, *70*(5), 1082–1097.

Cowan, N., Nugent, L. D., Elliot, E. M., & Saults, J. S. (2000). Persistence of memory for ingnored lists of digits: Areas of developmental consistency and change. *Journal of Experimental Child Psychology*, *76*, 151–172.

Cox, B. D., Ornstein, P. A., Naus, M. J., Maxfield, D., & Zimler, J. (1989). Children's concurrent use of rehearsal and organizational strategies. *Developmental Psychology*, *25*, 619–627.

Croy, C. D., & Novins, D. K. (2005). Methods for addressing missing data in psychiatric and developmental research. *Journal of American Academy of Child & Adolescent Psychiatry*, *44*, 1230–1240.

De Alwis, D., Myerson, J., Hersey, T., & Hale, S. (2009). Children's higher cognitive abilities and the development of secondary memory. *Psychonomic Bulletin & Review*, *16*, 925–930.

De Bruin, E. E. J., Van der Steene, R. G., Van Haasen, P. P., Coetsier, P., Pijl, Y. L., Spelders-Claes, R., et al. (1986). *Wechsler Intelligence Scale for Children (WISC-R)*. Lisse, The Netherlands: Swets & Zeitlinger.

DeMarie, D., & Ferron, J. (2003). Capacity, strategies, and metamemory: Test of a three-factor model of memory development. *Journal of Experimental Child Psychology*, *84*, 167–193.

Forrester, G., & Geffen, G. (1991). Performance measures of 7- to 15-year-old children on the Auditory Verbal Learning Test. *The Clinical Neuropsychologist*, *5*, 345–359.

Gaultney, J. F., Bjorklund, D. F., & Goldstein, D. (1996). To be young, gifted and strategic: Advantages for memory performance. *Journal of Experimental Child Psychology*, *61*, 33–66.

Glosser, G., Gallo, J. L., Clark, C., & Grossman, M. (2002). Memory encoding and retrieval in frontotemporal dementia and Alzheimer's disease. *Neuropsychology*, *16*, 190–196.

Huizingh, E. (2002). *Inleiding SPSS 11 voor Windows*. Schoonhoven, The Netherlands: Academic Service.

Kofman, O., Larson, J. G., & Mostofsky, S. H. (2008). A novel task for examining strategic planning: Evidence for impairment in children with ADHD. *Journal of Clinical and Experimental Neuropsychology*, *30*, 261–271.

Korkman, M., Kemp, S. L., & Kirk, U. (2001). Effects of age on neurocognitive measures of children ages 5 to 12: A cross-sectional study on 800 children from the United States. *Developmental Neuropsychology*, *20*(1), 331–354.

Kramer, J. H., Delis, D. C., Kaplan, E., O'Donnel, L., & Prifitera, A. (1997). Developmental sex differences in verbal learning. *Neuropsychology*, *11*, 577–584.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York, NY: Oxford University Press.

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment* (4th ed.). New York, NY: Oxford University Press.

Linschoten, J. (1963). *De La Court frequentietellingen van Nederlandse woorden*. [De La Court frequency countings of Dutch words] Utrecht, The Netherlands: University of Utrecht, Psychologisch Laboratorium.

Meijs, C. J. C. (2008). *Verbal learning in school-aged children and the influence of child-related factors, test-related factors, and natural context*. Maastricht, the Netherlands: Maastricht University.

Meijs, C. J. C., Hurks, P. P. M., Kalff, A. C., Slaats-Willemse, D. I. E., Rozendaal, N., & Jolles, J. (2009). Differential development of learning strategies on a Pictorial Verbal Learning Test (PVLT) in primary school children. *Child Neuropsychology*, *15*, 247–261.

O'Sullivan, J. T. (1996). Children's metamemory about the influence of conceptual relations on recall. *Journal of Experimental Child Psychology*, *62*, 1–29.

Ornstein, P. A. (1999). Comments: Towards an understanding of the development of memory. In F. E. Weinert & W. Schneider (Eds.), *Individual development from 3 to 12: Findings from the Munich Longitudinal Study* (pp. 94–105). Cambridge, England, UK: Cambridge University Press.

Saan, R. J., & Deelman, B. G. (1996). Nieuwe 15-woorden test A en B (15WT-A en 15WT-B). In A. Bouma, J. Mulder, & J. Lindeboom (Eds.), *Neuropsychologische Diagnostiek: Handboek*. [Neuropsychological Diagnostics: Handbook] (pp. 13–28). Amsterdam, the Netherlands: Swets & Zeitlinger.

Schlagmuller, M., & Schneider, W. (2002). The development of organizational strategies in children: Evidence from a microgenetic longitudinal study. *Journal of Experimental Child Psychology*, *81*, 298–319.

Schmidt, M. (1996). *Handboek RAVLT*. Los Angeles, CA: Western Psychological Services.

Schneider, W., Knopf, M., & Stefanek, J. (2002). The development of verbal memory in childhood and adolescence: Findings from the Munich Longitudinal Study. *Journal of Educational Psychology*, *94*, 751–761.

Schneider, W., Kron-Sperl, V., & Hunnerkopf, M. (2009). The development of young children's memory strategies: Evidence from the Wurzburg Longitudinal Memory Study. *European Journal of Developmental Psychology*, *6*, 70–99.

Schwenck, C., Bjorklund, D. F., & Schneider, W. (2009). Developmental and individual differences in young children's use and maintenance of a selective memory strategy. *Developmental Psychology*, *45*, 1034–1050.

Siegler, R. S., & Alibali, M. W. (2005). *Children's thinking* (4th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*(2), 428–444.

Sternberg, R. J., & Tulving, E. (1977). The measurement of subjective organization in free recall. *Psychological Bulletin*, *84*, 539–566.

Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. *Psychological Review*, *69*, 344–354.

Vakil, E., Greenstein, Y., & Blachstein, H. (2010). Normative data for composite scores for children and adults derived from the Rey Auditory Verbal Learning Test. *The Clinical Neuropsychologist*, *24*, 662–677.

Van der Elst, W., Van Boxtel, M. P., Van Breukelen, G. J., & Jolles, J. (2005). Rey's Verbal Learning test: Normative data for 1,855 healthy participants aged 24–81 years and the influence of age, sex, education, and mode of presentation. *Journal of the International Neuropsychological Society*, *11*, 290–302.

Van Loon-Vervoorn, W. A. (1989). *Voorstelbaarheidswaarden van Nederlandse woorden*. [Imageability ratings of Dutch words] Lisse, The Netherlands: Swets & Zeitlinger.